DOCUMENT RESUME

ED 409 713 FL 024 643

AUTHOR	Gibson, Bob
TITLE	Talking the Test: Using Verbal Report Data in Looking at the Processing of Cloze Tasks.
REPORT NO	ISSN-0959-2253
PUB DATE	97
NOTE	11p.; For journal issue as a whole, see FL 024 639. For other articles in this issue, see FL 024 640-646.
PUB TYPE	Journal Articles (080) -- Reports - Research (143)
JOURNAL CIT	Edinburgh Working Papers in Applied Linguistics; n8 p54-62 1997
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Cloze Procedure; Cognitive Processes; *Data Collection; Foreign Countries; *Language Processing; Language Research; *Language Tests; Linguistic Theory; *Protocol Analysis; *Research Methodology; Testing

ABSTRACT

The use of verbal report procedures as a research tool for gaining insight into the language learning process is discussed. Specifically, having second language students complete think-aloud protocols when they take cloze tests can provide useful information about what is being measured and how it has been learned. Use of such introspective information from the learner has regained credence among researchers, particularly when it is gathered during and not after the testing occurs. Real-time think-aloud reporting generates the largest amount of raw data concerning cognition, while retrospective protocols carry the risk that the learner is reconstructing, not recalling, cognitive processes he/she used. The greatest drawback of concurrent verbal reporting is the volume of data it provides; however, formats designed to limit the volume of information may also limit its quality. An alternative to the conventional think-aloud procedure for use with cloze tests is proposed. In this "annotated cloze" procedure the learner writes notes concerning the task as he/she completes it. While the annotated cloze provides somewhat less information than some other methods, the quality of the information obtained is considered to be better. Contains 14 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TALKING THE TEST: USING VERBAL REPORT DATA IN LOOKING AT THE PROCESSING OF CLOZE TASKS

Bob Gibson (DAL)

ED 409 713

FL024643

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Brian Parkinson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TALKING THE TEST: USING VERBAL REPORT DATA IN LOOKING AT THE PROCESSING OF CLOZE TASKS

Bob Gibson (DAL)

Abstract

This paper discusses the use of verbal report procedures as research tools, with reference to the writer's own investigation of the cognitive processes of second language cloze test-takers. The credibility of introspective data, and issues of informant training, language of reporting, and interviewer behaviour are considered. An alternative procedure called annotated cloze (AC) is also described and its pros and cons are discussed.

1 Introduction

In this paper I look at verbal reports as a source of insight into 'invisible' cognitive processes. I outline the more common verbal report procedures and discuss various aspects of their use in gathering data, as well as their application (and that of an alternative data-gathering procedure) to my own study of cloze test-taking behaviours. It is hoped that this paper will be of practical value to anyone contemplating the use of verbal report data in her or his own research.

My motivation for adding yet another study of cloze to the pile was simply that I wanted to know what this procedure - used in virtually every institution that I have taught in - might be measuring. Previous studies of cloze have tended to focus on the extent to which test takers successfully recover the words deleted from the passage - in other words, their score. While measures of cloze success tell us how well a given test has discriminated among those to whom it was administered, they reveal little about the processes actually involved in arriving at the successful recovery of deleted items. At the level of everyday test administration this may seem unimportant, but it has been argued (Bachman 1990) that before we can evaluate a test's validity we must have some insight into the construct the test is held to measure. Despite the enormous number of statistical analyses of the end-products of cloze test-taking, the question of what aspect(s) of linguistic proficiency cloze measures is (J. C. Alderson, personal communication) still very much open.

What all variants of cloze involve is the deletion, on a pseudo-random or rational basis, of textual items - classically individual words. Cloze has been variously identified as a measure of textual readability, an index of reading comprehension, and a measure of overall language proficiency in the target language. My hope is that cloze test-taker's 'real time' reports of their own task processing may provide some insight into the demands cloze makes on the individual.

2 The rehabilitation of introspective data

Verbal reporting is the most widely used means of eliciting introspective data - that information which informants provide about their own internal mental processes or states. I will not outline here the history of verbal report procedure as an investigative tool, for this is adequately reviewed in Matsumoto (1993) and Pressley and Afflerbach (1995). The main point is that, after several decades of rejection on the ground that they were insufficiently empirical, introspective procedures are once again widely accepted as legitimate data-gathering tools.

A key event in the rehabilitation of the notion that individuals can provide useful and valid information about their own mental operations was the publication of Ericsson and Simon (1984), which argued the case for employing verbal report procedures, in particular that known as 'think-

aloud', to elicit data which cannot readily be accessed by other means. Ericsson and Simon's model of verbal reporting holds that concurrent verbal report tasks such as think-aloud do not significantly interfere with the cognitive processes being reported. Their argument has been widely (and at times perhaps a little uncritically) adopted as a rationale for the use of verbal report procedures in a wide range of investigative contexts.

The crux of Ericsson and Simon's thesis is their claim that individuals can reliably report the currently active ('heeded') contents of their short-term memory[1], and that even recent cognitive operations, such as the previous steps in a problem-solving process, may be reliably recalled. Given the right stimulus, then, informants can provide valid information about their own mental processes while engaged in problem-solving tasks such as reading difficult L1 text, processing foreign language data etc.

Nisbett and Wilson's 1977 counter-argument to the use of verbal report data - that much of an individual's cognitive effort takes place below the level of conscious attention and is therefore unavailable for reporting - is to some extent deflected by Ericsson and Simon's careful delimitation of the kinds of internal data which may in their view be externalised: the most reliable verbal reports reflect just what is currently active (that is, being consciously attended) in short term memory at the moment of reporting. One corollary of this is that descriptions or explanations of cognitive processes are inherently less credible than 'raw' process data itself. By way of illustration, here (from my own data) are extracts from a verbal report by an individual attempting to work out the meaning of 'sukkari' in the Japanese sentence: 'Moo sensei no shigoto ni sukkari nareta' ('I've now got used to teaching.') In the first segment, the informant's verbal report looks very much like an unfiltered, moment-by-moment externalisation of his flow of thought:

> *umm* moo *now?* sensei *okay teacher ah ... teacher* no *teacher's work okay* sensei no shigoto ni *for? in?* sukkari ... sukkari nan desu ka? (*what's* sukkari?)

Data such as these are, Ericsson and Simon suggest, highly credible by virtue of this very lack of analysis. The obvious metacognitive refinement reflected in the following segment, however, would in their view render its data rather less convincing:

> sukkari sukkari *umm don't know this one maybe it's from* suki *to like hmm it must be an adjective because of the ending eh maybe I've come to like working as a teacher? or teaching has become likeable? it has a positive feel to it anyway (laughs)*

As the above extracts suggest, most of my informants' verbal reports comprised a fluent mixture of unmediated thought-fragments and more or less considered analysis. If this blend of raw and refined cognition is a mark of naturalistic task processing[2], we ought perhaps to be cautious about trying to sift out supposedly suspect metacognitions from the flow of verbal data.

Furthermore, the data I have collected suggest that for some informants analysis and hypothesis-making, as in the second extract above, represent a definite interruption to the processing task: textual processing seems to stop and to be taken up again at a non-contiguous point. Other informants, however, appear to be able to superimpose a higher level of analysis at minimal cost to their performance of the processing task. In the light of this, it may be difficult to view raw and refined data in the same terms for every verbal report subject.

The caveats above applies less strongly to retrospective verbal reports (see below) for here the danger of informants (re)constructing - rather than recalling - their cognitive processing is almost certainly greater. Ericsson and Simon note that an informant, asked why she phoned her mother the night before, will commonly offer some reason for the action, even if she had no conscious motivation for

it at the time. Questions about *why* a cognitive event occurred are seen as particularly likely to invoke reconstruction, and must therefore be used with caution.

2.1 What verbal report procedures involve

Verbal report data-elicitation procedures take a variety of forms. Informants can be asked to verbalise about their cognitive operations while carrying out a task (concurrent verbal report), or after the task has been completed (retrospective verbal report). It is worth pointing out here that verbal report procedures need not elicit only oral data: informants may also (as in Bailey's 1980 use of diary entries) report in writing on their cognitive behaviour, or (Cohen and Cavalcanti 1990) be asked to fill out 'open-ended' questionnaires.

Research using concurrent verbal reports has frequently combined some form of think-aloud task with a post-task interview. Think-aloud procedure typically requires the informant to keep up a more or less uninterrupted stream of verbal data while engaged in a particular task, ideally externalising as much of her task-related cognitive processing as she can, and it has been used to investigate general reading strategies (Block 1986), lexical inferencing, or the guessing of unknown vocabulary meaning (Haastrup 1991), translation (Krings 1988) and C-Test-taking processes (Feldmann and Stemmer 1987). With the possible exception of Block 1986, all of these studies used think-aloud to capture informants' cognitive processing more or less in real-time. In Block's study informants interrupted their processing at pre-set times in order to report on what had just occurred, which might better be seen as 'immediate retrospection'.

Real-time think-aloud reporting generates the largest volume of raw cognition. In part directed at the informant herself, and hence often difficult for another to follow, this data may require significant clarification. It is axiomatic that this should take place while the informant's task processing is still fresh in her mind, so that any post-task interview should follow as swiftly as possible. This leads to something of a dilemma: the researcher needs time to identify points of interest within the data, and yet the more time spent in doing so, the greater the risk that the informant's memory of her task-processing will have faded and the greater the chance that she will, however innocently, reconstruct rather than recall this when interviewed. The procedures I recommend in the appendix may ease the task of noting points requiring later attention, thus allowing the earliest possible retrospective interview.

2.2 Problems in the use of verbal report data

Perhaps the most serious drawback in using verbal report procedures - in particular concurrent reporting - is the very large volume of data that can result. The audio tape think-aloud protocols that I have gathered of informants' completing 40-item second language cloze passages provide a rich record of how they went about the task, but the transcript of a single tape (the average duration of recordings was just over 25 minutes) can easily run to twelve or fifteen pages of typescript. Moreover, the task of transcribing concurrent verbal report protocols is extremely demanding in terms of concentration as well as time: substantial stretches of a protocol may be acoustically very unclear, and several close listenings may be required to adequately decipher the content.

Some verbal report formats have the effect of limiting the amount of data elicited, or maximising the clarity of what is reported. Block's (1986) investigation of reading strategies, in which informants reported only at the end of each sentence in the stimulus passage is one example. Haastrup's (1991) investigation of lexical inferencing (i.e. guessing unknown vocabulary in a text) used a 'dyad' reporting format, in which two informants worked through an L2 passage more or less in tandem, sharing, discussing and evaluating information and inferences. Both discontinuous reporting as in Block's study and dyad reporting as in Haastrup's tend to generate data which is acoustically much clearer.

My own data-gathering trials, however, suggest that these formats may bring problems of their own. While sentence endings appear (Taft 1991) to be the loci of much of the higher level processing of text - such as the integration of discrete facts or notions into an overall text format - protocols generated by informants required to think aloud only at these points can seem rather disjointed and hard to follow. I also found it necessary to give reminders to 'keep talking' rather more frequently in this condition, and it may be that a format which allows verbal reporting to be switched off between sentence endings also makes it more likely that informants will forget to switch it on again.

Nor was dyad-condition reporting without drawbacks. Paired informants were observed to avoid reporting information which they thought their partner also had access to, and often failed to mention, inferences which they would also draw. Post-task interviews with (individual) informants elicited information and inferences which the informant claimed to have used in recovering deleted items, but which she had not reported during the paired task; these lacunae were often justified on the ground that there was no reason to tell a partner what she already knew[5]. Given the semi-conversational nature of dyad reporting tasks, it is not surprising that they appear to obey the Gricean maxim of economy.

One further limitation of the dyad format is that certain processing behaviours quite common in the solo think-aloud protocols I collected were more or less absent from those elicited from paired informants. These included the verbatim reading aloud of (often fairly long) stretches of the cloze passage, the 'sounding out' of alternative filler words, and the temporary substitution of nonsense words or sounds for deleted items. The range of processing behaviours open to individual informants, then, may not be adequately reflected in dyad reporting. In the last analysis, test-taking is normally a solitary activity, and the nature of the task is perhaps best maintained by having informants report individually.

2.3 Other aspects of data-elicitation

The choice of a reporting format is accompanied by other procedural decisions. How should the task passage - or other stimulus - be selected? Should informants be offered any training in verbal reporting? In which language should informants report?

2.3.1 Setting a task

In my own data-collection, a fair amount of trialling was required in order to find a productive elicitation text. While I could identify no clear relationship between a passage's difficulty, as measured by readability formulae, and the amount of verbal reporting it stimulated[6], it soon became clear that the passage had to be challenging enough to force informants to use a representative sample of their problem-solving abilities. Where the level of difficulty is low, (in cloze, for example, when the passage contains many deletions recoverable via syntactic knowledge alone), there may be very little in the way of conscious processing to be externalised. It may, then, be better to err on the side of passage difficulty. Naturally, the topic of the stimulus passage must be compatible with the assumed background knowledge of one's informants. While there may be marked discrepancies between the background knowledge one assumes will be available to a given group of informants and what they actually possess, it is hard to see how prior knowledge can be formally assessed without affecting the authenticity of the subsequent data-elicitation task. Caution in selecting the passage topic may be advisable.

2.3.2 Informant training

The question of whether to train one's informants in verbal reporting merits careful thought. On the one hand, few if any informants will have been asked to perform a verbal report task before, and they will typically be very unsure of what is expected of them. Clear instructions are thus vital, and some

BEST COPY AVAILABLE

sort of orientation session may be helpful. There is, however, (Haastrup 1991) a real risk of bias if informants are induced to alter their spontaneous processing behaviour to match that modelled in pre-task training; the very novelty for most informants of the think-aloud task may well magnify this danger. A workable compromise is to have potential informants attempt, without any external modelling, a short verbal report task broadly similar to that used in the actual data collection. During this phase informants are encouraged to report in more detail, but are given no guidance in how to go about this; any resulting changes in their task-processing behaviour will then at least be internally generated.

Pre-task orientation of this type can also help to minimise what I call the cold start effect, whereby informants attempting a think-aloud task for the first time may be unable to report in much detail, and only warm up, as it were, some way into the task. An orientation session may also allow the researcher to exclude any informants who produce little or no think-aloud data. This idea of sidelining potential informants may seem like poor empirical practice, but it can be defended on two grounds. Firstly, the risk of bias is minimised by the fact that the criterion for exclusion is the failure to provide data, rather than the kind of data that is provided. Furthermore, even if non-reporters are able to retrospect in detail about their task processing, there will be no concurrent data against which to evaluate these retrospections. Such triangulation of data from different sources is a desirable (Fielding and Fielding 1986) aid to the validation of research findings.

2.3.3 Language of reporting

My informants to date have mainly been speakers of German, simply because this is the second language in which I am most fluent. In a number of earlier studies using verbal report data (e.g. Block 1986), informants were required to verbalise in a second language - i.e. the first language of the researcher. Setting a second language as the language-of-reporting, of course, carries the rather obvious risk that those cognitive operations which the informant cannot adequately report in her L2 will simply go unrecorded. In my own data-collection informants were allowed to verbalise in German or in English as they preferred, but it was made clear that L1 reporting was entirely acceptable. The main argument for L1 reporting is, of course, that this imposes the least additional burden - and hence bias - on the processing of the task itself. In the event, most of my informants reported very largely in their L1, and only a few who showed a high level of confidence in their English ability (a confidence which was not always reflected in their cloze-task scores) chose to report in English; even these never avoided the use of their L1 entirely. The use of L1 reporting is not entirely straightforward, however: there exists, for instance, the possibility that my encouragement of L1 reporting may be partly responsible for the frequency with which my informants translated often quite long stretches of the task passage into their L1.

2.4 The interaction of researcher and informant

Arguably, the researcher should interact with informants as little as possible during the data-elicitation task - yet this may be unavoidable. However clearly the verbal report task has been explained or even modelled, an informant may still have questions or go astray. In my own data-gathering sessions, for example, I quite frequently had to re-emphasise (although these things appeared to have been clearly understood beforehand) that only one filler-word was allowed per cloze deletion, and that every deletion required a word. Reminders to verbalise may also be needed: perhaps the least interruptive technique is simply to point to a sign that reads 'Please keep talking'. Eye-contact (see Appendix) with informants should be minimised.

2.5 Post-task interviewing

As the think-aloud task tends to be quite tiring for informant and researcher alike, a short break may be required before any post-task interview. This interval also allows the researcher to very quickly

review her notes, and to prepare questions about any aspects of the informant's processing which were unclear, or which otherwise merit following up.

Post-task interviews may be more or less structured according to the researcher's purposes, but at least some questions may be prepared in advance on the basis of processing behaviours observed in trial verbal report sessions. Some degree of structure is of course vital if cross-comparison among informants is intended, but retrospective interviews work best when the informant feels free to make spontaneous comments and does not merely respond to questions. In my own data-gathering, informants' spontaneous comments have generated as much insight as their responses to my prepared prompts.

An informant's audio recording can be used as necessary during the post-task interview, in the role of aide-memoire. In my own experience, however, it is better to let the informant retrospect about a given aspect of the task as far as possible without cueing from the recording; only where little or no information is forthcoming should the tape be used as a memory support. The researcher may of course wish to know more about a particular segment of the verbal report, in which case replaying of the appropriate extract may be unavoidable. It is unrealistic to expect informants to recall (rather than reconstruct) details of their processing without this support.

Replaying of the recording can also have a control function, in that inconsistencies may be found between an informant's retrospective remarks and the content of her on-task verbal report. To cite one example, an informant who claimed she had not translated parts of the stimulus passage into German was heard to do so quite extensively on the recording[7]. By and large - although minor discrepancies between on-task and post-task reports are not uncommon - informants' retrospections tend to be reliable if the interval between task and retrospective interview is short. Where an informant's assertions about her task processing conflict markedly with the data on her tape, it may be best simply to discard that individual's report completely. There is little to be gained, in my view, from embarrassing the informant with evidence of her own misapprehensions - especially if (the 'snowball' sampling procedure) one is relying on her to help find other volunteer subjects. Clearly, however, there are risks in simply taking informants' retrospections at face value, and confirmatory evidence from other data sources should always be sought.

3 An alternative to think-aloud procedure?

To sum up so far, concurrent think-aloud procedure provides arguably the richest and most direct reflection of cloze processing behaviours, but at very high cost in time and effort. Having used think-aloud in compiling a working taxonomy of cloze-processing strategies, I began to look for a less labour-intensive means of eliciting at least a subset of the same information from a larger group of informants. With this in mind, I revised an instrument which I had tentatively explored some years ago, and have labelled 'annotated cloze' (AC). Here the informant is required not to verbalise her processing, but to write it down in real time during the task. Cues within the text that aid recovery of deleted cloze items are underlined or circled, with arrows indicating the deletion(s) to which they were applied. Extratextual information aiding recovery is written alongside the appropriate deletion. Again, informants are encouraged to report in the L1 if they prefer, on the ground that this creates the least additional burden on processing.

3.1 Pros and cons of the AC task

The potential drawbacks of this procedure are obvious. For one thing, given that it takes much longer to write down an inference than to verbalise it, we might anticipate a greater loss of data than would be the case with oral reporting. Furthermore, AC provides no indication of how much time was spent on a given deletion - something which can be assessed fairly accurately from an oral protocol. AC, moreover, tells us very little about the relative difficulty informants experience with sub-tasks such as

individual cloze deletions. Finally, AC may largely be limited to recording the textual data and extratextual knowledge which informants employ in completing the task.

But the picture is not entirely dismal, for AC procedure also has several advantages over think-aloud. One positive aspect is that, needing no input from the researcher, AC data can be gathered from a number of informants simultaneously, or (provided instructions to informants are clear) at any geographical remove. The collation and initial analysis of AC data is also much less onerous: requiring no transcription stage, an informant's AC manuscript can be analysed much faster than any think-aloud recording.

Another benefit is that, although certainly less comprehensive than think-aloud data, the data elicited in AC tasks tends to be more focused. By this I mean that informants seem to edit out - consciously or unconsciously - those cognitions which were less central to task completion, and thus present their data in a rather more condensed and ordered fashion than would be the case in an oral verbal report protocol. Given the constraints imposed by a written form, this is hardly surprising. Still another positive aspect of AC is that informants have a ready overview of the information they have already volunteered, and are thus better able to expand upon or to amend this than would be practical in an oral report. AC manuscripts often contain such revisions[4] and clarifications, in fact, as well as comments ('Eh?', 'Difficult!') that reveal something of the informant's affective response to the task.

Finally, the information lost in substituting an AC task for oral think-aloud may be partially recovered via the use of a post-task questionnaire. Successful questionnaire design is notoriously problematic and the desirability, mentioned above, of encouraging informants to volunteer insights into their task-processing adds to the complexity of the design task. Draft questionnaires may usefully be trialled through one-on-one observation of, and discussion with, the informants tasked with completing them. As my own questionnaires are in the informants' L1, I have found it helpful to ask informants themselves to revise the questions for clarity, as well as to propose new questions. By asking informants to think aloud as they work through the post-task questionnaire, in the presence of myself or of a tape recorder, I have been able to trap valuable insights which might otherwise have been lost.

3.2 The credibility of AC data

If we follow Ericsson and Simon, the data provided by AC protocols are clearly more refined (and hence less credible) than the rawer data contained in think-aloud reports. As I suggest above, however, the data reported during the processing of text-based tasks seem to exist along a cline of mediation, ranging from bits of words up to considered analyses. Indeed, my own reading of published studies which justify their methodologies by appeal to Ericsson and Simon's model suggests that not a few of these incorporate data which a strict application of Ericsson and Simon's approach to think-aloud would disallow. There seems, then, to be scope for the individual researcher's discretion in delimiting her data.

To sum up, although the risk of information loss through the use of AC is a serious one, comparison of its products with those of think-aloud protocols should provide some insight into what the former procedure fails to record. On that basis, I hope to be able to refine the AC task and its accompanying questionnaire, and to apply these to the gathering of data from a larger number of informants than would be possible with an oral report procedure.

Notes

[4] Ericsson and Simon's construct of 'short term memory' appears to this writer to be congruent with the more recent term 'working memory' (cf. Taft 1991).

[2] As observation of one's own real-time processing of texts and other data may suggest is the case.

[3] A certain amount of friction between partners was in fact evident when one failed to observe this convention.

[4] This was true of both cloze and intact reading passages.

[5] The situation was less clear cut regarding another informant who claimed to have translated quite extensively, but whose recording contained little evidence of this.

[6] A key instruction to informants is that revisions must be added above the (struck through) original data; nothing may be erased.

References

Bachman, L. 1990. *Fundamental Considerations in Language Testing*. Oxford: OUP.

Bailey, K. 1980. 'An introspective study of an individual's language learning experience', in Scarcella, R. & S. Krashen (eds.) *Research in Second Language Acquisition: Selected Papers of the Los Angeles Second Language Research Forum*. Rowley, MA: Random House: 58-65.

Block, E. 1986. 'The comprehension strategies of second language readers'. *TESOL Quarterly* 20/3: 463-494.

Cohen, A. and M. Cavalcanti. 1990. 'Feedback on compositions: teacher and student verbal reports' in Kroll, B. (ed.) *Second Language Writing. Research Insights for the Classroom*. Cambridge: CUP: 155-177.

Ericsson, K. and H. Simon. 1984 (revised edition 1993) *Protocol Analysis. Verbal Reports as Data*. Cambridge, MA: MIT Press.

Feldman, U. and B. Stemmer. 1987. 'Thi_aloud a_ retrospective da_ in C-te_taking: diffe_languages -diff_learners - sa_approaches?' in Faerch, C. & G. Kasper. 1987. *Introspection in Second Language Research*. Clevedon, England: Multilingual Matters.

Fielding, N. and J. Fielding. 1986. *Linking Data* (*Qualitative Research Methods*, Volume 4). Beverley Hills: Sage.

Haastrup, K. 1987. 'Using thinking aloud and retrospection to uncover learners' lexical inferencing procedures', in Faerch & Kasper 1987: 197-212.

Haastrup, K. 1991. *Lexical Inferencing Procedures or Talking About Words*. Tübingen: Gunter Narr.

Krings, H. 1988. 'Blick in die 'Black Box': Eine Fallstudie zum Übersetzungsprozeß bei Berufsübersetzern', in Arntz, R. (ed.) *Textlinguistik und Fachsprache. Akten des Internationalen Übersetzungswissenschaftlichen AILA-Symposions*, Hildesheim, April 1987. Hildesheim: Olms: 393-412.

Matsumoto, K. 1993. 'Verbal report data and introspective methods in second language research'. *RELC Journal* 24/1: 32-60.

Nisbett, R. and T. Wilson. 1977. 'Telling more than we can know: verbal reports on mental processes'. *Psychological Review* 84: 231-259.

BEST COPY AVAILABLE

Pressley, M. and P. Afflerbach. 1995. *Verbal Protocols of Reading*. Hillsdale, N. J.: Erlbaum.

Taft, M. 1991. *Reading and the Mental Lexicon*. Hove: Erlbaum.

APPENDIX: THE MECHANICS OF GATHERING VERBAL REPORT DATA

The practical aspects of collecting verbal report data are quite sparsely discussed in the literature, so it may be useful to describe here the procedures which I have arrived at by trial and error.

When tape recording verbal report data, it is important to use a good quality microphone (the microphone built into the average cassette recorder may not be sensitive enough) and to place this close enough to the informant to pick up even low-volume speech. To minimise masking through the noise of rustling paper etc., the stimulus task should wherever possible be presented on a single page. In a dyad reporting task it is especially important to obviate the need to turn a page, as this is very likely to disrupt one informant's processing.

The cassette recorder itself should be placed close to the researcher so that (a) the informant is not tempted to rewind or otherwise interrupt the recording and (b) the researcher can easily view the tape-counter window. Experience suggests that using a clock or watch to keep track of informants' progress may lead them to think that is a factor in their performance. As this risks modifying informants' behaviour, it is of course better avoided. It is in fact less awkward to keep track of informants' processing solely by means of the tape counter as this makes it possible to note the precise location on the tape of interesting events, and to return to these during the post task interview. Accurate timings can of course be made at a later stage. Some cassette recorders have a very useful 'bookmark' or 'memo' function, which allows points of interest on the tape to be tagged simply by pressing a button. The tape can quickly and easily be rewound to these points, saving a surprising amount of time if the tape has to be referred to in post-task interviews.

Having the researcher busy herself with the tape-counter has the additional advantage of distancing her to some degree from the informant's task processing. My experience has been that solo informants in particular are prone to seek eye-contact with the researcher, and if this is established they may go on to seek confirmation of hypotheses or answers. Paired informants, of course, tend to seek this kind of support from each other.

A pitch control is another useful cassette recorder feature. This allows tape speed to be increased or slowed by (typically) up to 20%. This is most helpful when transcribing and/or coding the recorded verbal reports: the protocols of informants who verbalise clearly can often be processed at faster than normal playback speed, while the ability to slow the playback allows less clear verbal reports to be transcribed or analysed with minimum use of the rewind key. Note too that the task of transcribing extensive audio protocols is made markedly less unpleasant by the use of a tape player with a foot-controlled review function - especially if the duration of the segment reviewed can be set incrementally.

Finally, choose cassette tapes with enough capacity per side to capture an informant's on-task verbal report in full: changing tapes in mid session is an unwelcome intrusion for both parties.